License to evaluate: Preparing learning analytics dashboards for educational practice

DOI:
10.1145/3170358.3170421

Document status and date:
Published: 01/03/2018

Document Version:
Peer reviewed version

License to evaluate: Preparing learning analytics dashboards for educational practice

Ioana Jivet
Open University of the Netherlands
Heerlen, the Netherlands
ioana.jivet@ou.nl

Maren Scheffel
Open University of the Netherlands
Heerlen, the Netherlands
maren.scheffel@ou.nl

Marcus Specht
Open University of the Netherlands
Heerlen, the Netherlands
marcus.specht@ou.nl

Hendrik Drachsler
Goethe University Frankfurt/DIPF, Germany
Open University of the Netherlands
Heerlen, the Netherlands
drachsler@dipf.de

ABSTRACT

Learning analytics can bridge the gap between learning sciences and data analytics, leveraging the expertise of both fields in exploring the vast amount of data generated in online learning environments. A typical learning analytics intervention is the learning dashboard, a visualisation tool built with the purpose of empowering teachers and learners to make informed decisions about the learning process. Related work has investigated learning dashboards, yet none have explored the theoretical foundation that should inform the design and evaluation of such interventions. In this systematic literature review, we analyse the extent to which theories and models from learning sciences have been integrated into the development of learning dashboards aimed at learners. Our analysis revealed that very few dashboard evaluations take into account the educational concepts that were used as a theoretical foundation for their design. Furthermore, we report findings suggesting that comparison with peers, a common reference frame for contextualising information on learning analytics dashboards, was not perceived positively by all learners. We summarise the insights gathered through our literature review in a set of recommendations for the design and evaluation of learning analytics dashboards for learners.

CCS CONCEPTS

• **Human-centered computing → Information visualization**; • **Information systems → Data analytics**; • **Applied computing → *E-learning***;

KEYWORDS

learning dashboards, learning theory, learning analytics, systematic review, learning science, social comparison, competition, evaluation

ACM Reference format:
Ioana Jivet, Maren Scheffel, Marcus Specht, and Hendrik Drachsler. 2018. License to evaluate: Preparing learning analytics dashboards for educational practice. In *Proceedings of International Conference on Learning Analytics and Knowledge, Sydney, NSW, Australia, March 7–9, 2018 (LAK '18),* 10 pages.
https://doi.org/10.1145/3170358.3170421

LAK '18, March 7–9, 2018, Sydney, NSW, Australia

ACM ISBN 978-1-4503-6400-3/18/03.
https://doi.org/10.1145/3170358.3170421

1 INTRODUCTION

Learning Analytics (LA) aims to exploit the potential of the increasingly large amounts of data describing interaction data, personal data and academic information generated by the widespread use of online learning environments [22]. LA is a relatively new field established in 2011 and defined as "the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs" [55].

Ruipérez-Valiente et al. [49] identified two main approaches of using LA to make sense of the vast amount of learning data: (i) building systems that rely on automatically processing the data, i.e., automatic actuators, such as intelligent tutoring systems, recommenders or adaptive systems, and (ii) reporting data directly to stakeholders. Recently, there has been a shift from mining learning data to support automated interventions towards directly reporting the data, usually in a visual form, thus empowering learners and teachers to better leverage human judgement [5]. This trend is visible in the increasing amount of research efforts invested into building learning dashboards and other visual representations of learning data with the purpose of improving student learning and performance in a variety of learning contexts.

As defined by Schwendimann et al. [53], learning analytics dashboards are "single displays that aggregate different indicators about learner(s), learning process(es) and/or learning context(s) into one or multiple visualisations". Although so far dashboards have been developed for a broad range of stakeholders, including learners, teachers, researchers or administrative staff, LA dashboards have the potential to be used as powerful metacognitive tools for learners, triggering them to reflect and examine their learning behaviour and learning outcomes [14, 20]. However, a recent literature survey [53] showed that the majority of dashboards supports teachers and not as much attention has been put into developing dashboards for learners and studying the effects such tools have on learners [33].

Learning analytics lies at the intersection between technology and learning science. Greller & Drachsler [27] envisioned learning analytics as an educational approach guided by pedagogy and not the other way around. However, there is a strong emphasis on the "analytics", i.e., computation of the data and creation of predictive models, and not so much on the "learning", i.e., applying and researching LA in the learning context where student outcomes can be improved [25]. This is seen as a critical shortcoming because "learning analytics is about learning" [25, 56]. Several other authors support the claim that there is a need for developing learning analytics tools based on learning theories as well as integrating them into the instructional design [39, 40].

As a first step towards building useful dashboards for learners, we need to understand how to consider learning sciences throughout the whole life-cycle of learning analytics dashboards: in the design and during the evaluation as well as during the pedagogical use of learning dashboards. Suthers & Verbert [56] argue that learning analytics research should be explicit about the theory or conception of learning underlying the work and manifest this conception in the presented work. Following their position, we wish to investigate which educational concepts form the theoretical foundation for the development of learning dashboards aimed at learners and how they are used in the design and evaluation of such dashboards.

Learning analytics dashboards literature has been reviewed in previous works. Verbert et al. [61] defined a conceptual framework for analysing learning analytics applications and examined 15 dashboards based on the target user group, the displayed data and the focus of the evaluation. A follow-up review [62] extended this analysis to 24 dashboards, analysing the context in which the dashboards had been deployed, the data sources, the devices used and the evaluation methodology. Yoo et al. [66] proposed an evaluative tool for learning analytics dashboards based on Few's principles of dashboard design [24] and Kirkpatrick's four-level evaluation model [32]. Using this tool, they reviewed the design and evaluation of 10 educational dashboards for teachers and students. A recent systematic review by Schwendimann et al. [53] of 55 dashboards looked at the context in which dashboards had been deployed, their purpose, the displayed indicators, the technologies used, the maturity of the evaluation and open issues.

All these works reviewed learning analytics dashboards, regardless of their target users. We narrowed the scope of this review to LA dashboards aimed at learners, addressing one of the challenges of learning analytics identified by Ferguson [22]: lack of focus on the perspectives of learners. A closely related work to this paper was published by Bodily & Verbert [9]. They have analysed student-facing LA systems, including dashboards, educational recommender systems, EDM systems, ITS and automated feedback systems. The systems were reviewed based on their functionality, the data sources, the design, the effects perceived by learners and their actual effects.

The learning-theory foundations of game-based learning have already been investigated in order to underpin educational computer game design [65], but none of the previous learning dashboard reviews addressed the connection to learning sciences. Moreover, Schwendimann et al. [53] and Bodily & Verbert [10] provided recommendations for the design of learner dashboards, but neither suggested the use of educational concepts as a basis for the design or evaluation of the dashboards, despite recommendations made by other authors [40, 56]. Through this systematic literature review, we aim to address this gap by investigating the relation between educational concepts and the goals and evaluation of learning dashboards. Dashboard evaluation was previously examined by looking at the experimental setting, learners' perceived usability and effects regarding performance, skill or behaviour change. However, in this study, we will focus on the criteria used for dashboard evaluation, how these criteria relate to the goal of the dashboard, what type of data was used for the evaluation and how learners perceived the framing of the information to ease sense-making.

1.1 Previous work

In a previous literature review study [30], we investigated the relation between learning sciences and learning analytics by looking into which educational concepts inform the design of learning analytics dashboards aimed at learners. Our main findings show the most common foundation for LA dashboard design is self-regulated learning theory, frequently used to motivate dashboard goals related to supporting awareness and triggering reflection. However, current dashboards offer some support only for the performance phase of the SRL cycle as defined by Zimmerman et al. [68], as their primary goal is to support awareness. Nevertheless, just making learners aware is not enough as awareness does not imply that actions are being taken. We argue that dashboards should have a broader purpose, using awareness and reflection as means to improve cognitive, behavioural or emotional competences. Additionally, since very few dashboards are adequately integrated into the learning environment or the learning design, learners miss support in the other two phases, i.e., forethought and self-reflection. Furthermore, the majority of dashboards use comparison with peers as a representative frame of reference for evaluating their performance. Frames of references are anchor points that learners can use in order to make sense of and evaluate the data displayed on the learning dashboard [64]. Thus, there is a strong emphasis on comparison and competition with peers, although research in educational sciences identifies different sources of motivation for learners, i.e., performance and mastery achievement goal orientations [47].

1.2 Approach

Throughout this literature review, we build upon this previous study and continue to explore how the educational concepts and learning theories are used through the whole development cycle of the dashboard, focusing on how they are integrated into the evaluation of learning dashboards. More specifically, we look at how the evaluation was conducted, what was evaluated and how it relates to the goal of the dashboard and the educational concepts cited. Therefore, our study is guided by the following research questions:

(1) How does the evaluation relate to the purpose of the dashboard?
(2) How are the educational concepts considered in the evaluation?
(3) How do learners perceive different frames of reference?

In the next section, we will briefly describe the methodology we used for identifying relevant literature. In Section 3 we present the results of our analysis, while in Section 4 we discuss the key insights of our review. Lastly, Section 5 provides concluding remarks.

2 METHODOLOGY

To answer the research questions described in the previous section, we conducted a systematic review following the PRISMA statement [43]. An informative literature search preceded the systematic review to get an overall picture of the field. For our search, we selected the following databases as they contain relevant literature for the field of learning analytics: ACM Digital Library, IEEE Xplore, SpringerLink, Science Direct, Wiley Online Library, Web of Science and EBSCOhost. Additionally, we included Google Scholar to cover any other sources, limiting the number of retrieved results to 200. We queried the selected databases with the following search terms: *"learning analytics" AND (visualization OR visualisation OR dashboard OR widget).* The first term focuses the search on the field of *learning analytics*, while the second part of the query is meant to cover different terminologies used for this type of intervention, addressing one of the limitations identified in [53]. Although the scope of this review is limited to visualisations that have learners as end-users, it was not possible to articulate this criterion in relevant search terms. Therefore, we built a query that retrieved all dashboards, regardless of their target end-users, and removed the ones that fell out of our scope in a later phase.

The queries were run on February 20th, 2017, collecting 1439 hits. We further screened for relevance each of the results by examining the title and the abstract, removing articles that do not describe dashboards, widgets or visualisations aimed at learners or do not have the full text available in English. Once we removed any duplicates, the list of potential candidate papers was reduced to 212. Eleven papers that we came across during the preliminary check and fit the scope of our survey but did not appear in the search databases were also added to the set of papers to be further examined. Next, we accessed the full text of each of these 223 studies to assess whether they are eligible for our review.

We retained only the papers that: i) describe dashboards for learners, ii) describe fully developed dashboards or at least one iteration of its development, i.e., we excluded any theoretical or conceptual papers, essays or literature reviews, iii) explicitly mention educational concepts behind the goals, design or evaluation of the dashboard, iv) include an evaluation of the dashboard. We identified 97 papers that satisfied the first two criteria. Among these papers, only 50 explicitly mentioned theories, models or concepts borrowed from the field of learning sciences. Similarly, just 54 out of 97 papers identified through our search contained an evaluation of the introduced dashboard.

The focus of this study as well as that of the previous literature review published in [30] is set on the 28 papers that describe dashboards together with their empirical evaluation and rely on learning theories. As a final selection filter, we removed two papers that presented preliminary work on the same dashboard as the content of these papers was covered in later works. Finally, our literature survey included 26 papers that satisfied all our criteria.

3 RESULTS

We started our analysis by identifying which aspects each dashboard evaluation focused on and how the intended goal of the dashboard related to the focus of the evaluation. Next, we looked at how the different aspects were evaluated and how the educational concepts on which the design relied were taken into account in the evaluation. We closed our analysis by examining how learners perceive the use of different frames of reference.

3.1 Dashboard evaluation levels

Taking a closer look at the evaluation sections of each of the papers included in this study, we distinguished twelve criteria that were assessed during the evaluation of the dashboards. Following the same rationale used in [30] for clustering dashboard goals, we grouped these criteria into four levels based on the competence they aimed to affect in learners: metacognitive, cognitive, behavioural or emotional (see Table 1). A fifth level *EV5: Self-regulation* was added to account for papers that specifically evaluated and measured improvements in self-regulated learning, an ability that entails all four other competences [67]. Finally, the sixth level *EV6: Tool usability* was included to cover aspects related to the usability of the tool.

Table 1: Levels, criteria for evaluation and the papers in which they were used.

Level	Criterion	Freq.	Papers
EV1: Metacognitive	Understanding	11	[4, 15, 16, 18, 28, 34] [37, 42, 50, 51, 63]
	Agreement	3	[34, 37, 42]
	Impact on awareness and reflection	11	[15, 16, 18, 34, 37, 38] [42, 50, 51, 54, 58]
EV2: Cognitive	Impact on performance	12	[7, 16, 19, 28, 29, 31] [36, 38, 50, 51, 54, 57]
EV3: Behavioural	Impact on behaviour	17	[4, 7, 16, 18, 19, 28, 29] [34, 37, 38, 42, 50, 51] [54, 57, 58, 63]
	Usage of the system	7	[7, 13, 29, 31, 38, 50, 63]
EV4: Emotional	Impact on motivation	4	[16, 28, 38, 58]
	Impact on affect	2	[44, 63]
EV5: Self-regulation	Self-regulated learning	1	[57]
EV6: Tool usability	Satisfaction	2	[31, 38]
	Usability	8	[3, 12, 28, 31, 38, 50] [57, 58]
	Usefulness	13	[12, 13, 26, 28, 31, 34] [35, 37, 38, 42, 50, 58] [63]

The *EV1: Metacognitive level* groups aspects related to learners' knowledge, beliefs and reflection on their learning processes, strategies and their effectiveness. The three aspects included here are understanding of the information displayed on the dashboard, agreement with this information and the impact the dashboard has on learners' awareness and reflection. The *EV2: Cognitive level* contains aspects that evaluate learners' understanding and knowledge regarding the studied material and was operationalised through their performance and the quality of their learning outcomes. The impact of the dashboard on the learners' engagement, online social behaviour and help-seeking behaviour are grouped under the *EV3: Behavioural level.* We also included learners' usage of the learning environment and the dashboard itself under this level. The *EV4: Emotional level* consists of two aspects: impact on motivation and impact on affect. The final category, *EV6: Tool usability*, gathers aspects related to the tool itself, its acceptance, ease of use, usefulness, as well as learners' satisfaction.

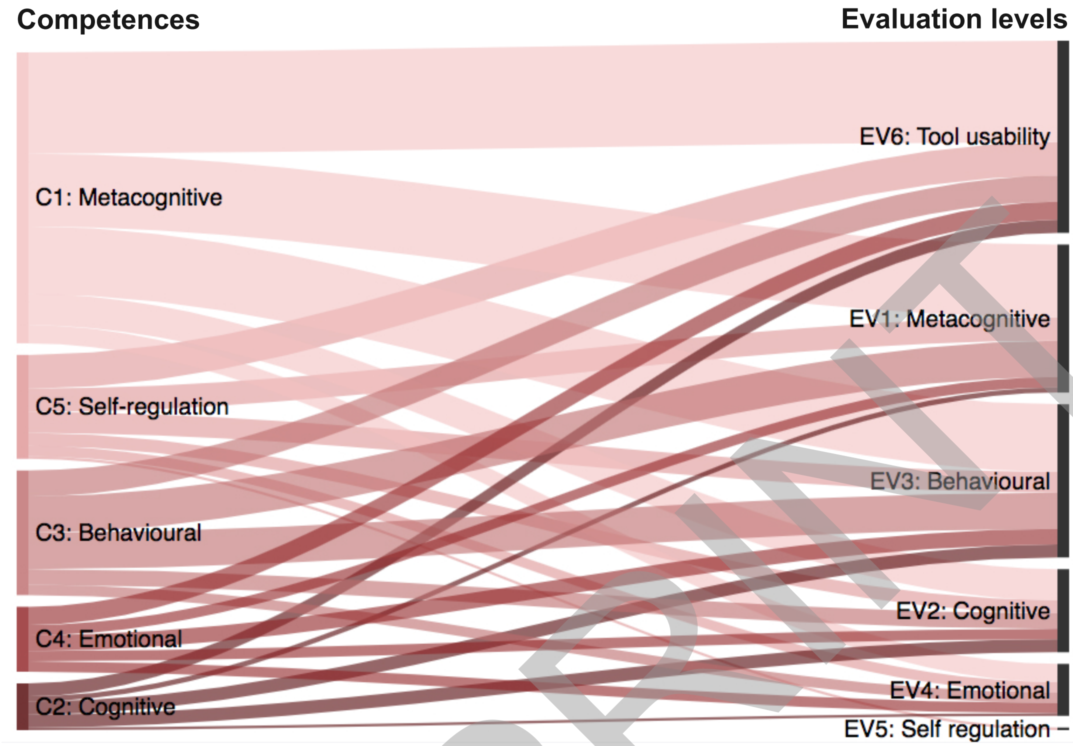

Figure 1: The competence level targeted by the dashboard included in the review in relation to the criteria assessed in the evaluation of the dashboard

Table 2 lists the papers that targeted the five competences with their dashboard and the papers that evaluated changes in these competences. The column "coverage" represents the percentage of papers that both targeted and evaluated each competence. The value of this indicator for the five categories shows that behavioural, cognitive and emotional levels were evaluated in most of the cases where the dashboard was designed to support this competence. However, a large percentage of dashboards that target metacognition and self-regulation were missing an evaluation. Surprisingly, several dashboards were evaluated also on dimensions that were not in line with the goal for which they were designed, especially on the cognitive and behavioural level. We hypothesise that this happens due to the availability of behavioural data extracted from trace logs and the assessment data used to estimate learners' performance.

Table 2: The list of papers that targeted and evaluated in their dashboard design each competence. Coverage represents the percentage of papers that both targeted and evaluated each competence.

Competence	Targeted in	Freq.	Evaluated in	Freq.	Coverage
C1: Metacognitive	[3, 4, 12, 13] [15, 18, 19, 26] [28, 31, 34, 35] [37, 38, 42, 44] [50, 51, 54, 57] [58, 63]	22	[4, 15, 16, 18] [28, 34, 37, 38] [42, 50, 51, 54] [58, 63]	14	59%
C2: Cognitive	[7, 19, 29, 31] [44, 50]	6	[7, 19, 28, 29] [31, 36, 38, 50] [51, 54, 57]	11	83%
C3: Behavioural	[4, 7, 18, 19] [28, 29, 34, 37] [38, 51, 58, 63]	12	[4, 7, 13, 16] [18, 19, 28, 29] [31, 34, 37, 38] [42, 50, 51, 54] [57, 58, 63]	19	100%
C4: Emotional	[28, 38, 44, 58] [54]	5	[16, 28, 38, 44] [58, 63]	6	80%
C5: Self-regulation	[12, 15, 19, 26] [28, 35, 36, 42] [44, 50, 57, 58] [63]	13	[57]	1	8%

Figure 1 illustrate the relation between the competence targeted by the dashboards and the evaluation levels. It is important to note that each paper may be represented several times, as most of the dashboards served various goals and were also evaluated according to multiple criteria. These results show that most of the evaluations focused on the tool's usability and the metacognitive and behavioural levels. Comparatively, very little attention was given to the effects on the cognitive, emotional and self-regulation levels. The tool's usability, usefulness and learners' satisfaction with the tool were evaluated in more than half of the analysed papers (14/26) and changes in learners' learning behaviour were discussed in 17 papers. Seven papers also looked into how frequent or how the dashboard was used. Surprising is the fact that only one paper directly evaluated improvement in self-regulated learning although 13 papers listed supporting self-regulated learning as one of their goals [30]. When it comes to the evaluation of the metacognitive competence, the majority of the papers included in this study aim to support awareness and reflection (20/26) [30], yet only half of those (11) assessed whether the dashboard had any impact on the learners' awareness and reflection. Furthermore, half of these papers examined whether learners can make sense of the information displayed on the dashboard, but only 3 out of these took a step further and questioned whether learners agreed with the data shown to them. We identified just seven papers that evaluated criteria from only one level, while the majority of papers considered two evaluation levels (5 papers), three levels (7), four levels (5) and five levels (2). None of the papers evaluated all six evaluation levels.

3.2 Dashboard evaluation data

We classified the data used in the evaluation of the dashboards into i) self-reported by learners, ii) tracked data that was automatically collected by the system, and iii) assessment data (see Table 3).

Table 3: Data used in the evaluation of dashboards.

Category	Data type	Frequency	Papers
Self-reported	Feedback survey	17	[3, 4, 12, 13, 15, 18, 26] [28, 31, 35, 37, 38, 51] [42, 50, 57, 58]
	Interview	3	[16, 50, 63]
	Focus group	2	[34, 58]
	Evaluation instrument	6	[7, 16, 28, 29, 54, 57]
Tracked	Resource use	8	[18, 19, 28, 37, 38, 51] [57, 63]
	Learning artefacts	7	[7, 19, 28, 29, 38, 51, 63]
	Dashboard use	9	[7, 13, 28, 29, 31, 44] [38, 50, 57]
Assessment	Grades	7	[19, 28, 31, 36, 51, 54, 57]

20 of the papers collected self-reported data in the form of surveys, interviews or focus groups to assess the dashboards in most of their aspects. Seven papers also used standardised instruments for measuring approaches to learning (in [29]), self-regulated learning skills (in [57]), time management skills (in [57]), achievement goal orientations (in [7, 28]), learning motivations (in [16]) and learning power (in [54]) through validated measuring instruments from learning sciences in order to assess the impact of the dashboard on behaviour or affect. These instruments include the 3x2 AGQ instrument to determine students' Achievement Goal Orientations [21], the Revised Two-Factor Study Process Questionnaire (R-SPQ-2F) to determine students' approaches to learning [8], the Online Self-Regulated Learning Questionnaire (OSLQ) for measuring self-regulation in online and blended learning environments [6], the Validity and Reliability of Time Management Questionnaire (VRTMQ) to evaluate how learners manage their time [1] and the Motivated Strategies for Learning Questionnaire (MSLQ) to establish learners' motivations relating to the subject of the course [48]. 14 dashboard evaluations used trace logs. These data were automatically collected by the platform and were used for assessing whether there had been any changes in learners' use of resources or the quality of their learning artefacts. For the purpose of this work, we grouped indicators describing objects produced by learners during the learning process, e.g., discussion forum posts, as well as the content and efficiency of their assessments under the term "learning artefacts", e.g. quiz question attempt, learning gains, quiz question efficiency or success rate. Nine papers collected data that describes the use of the dashboard itself. Assessment data was used in seven papers to determine whether using the dashboard had any effect on students' performance measured through grades or graduation rate.

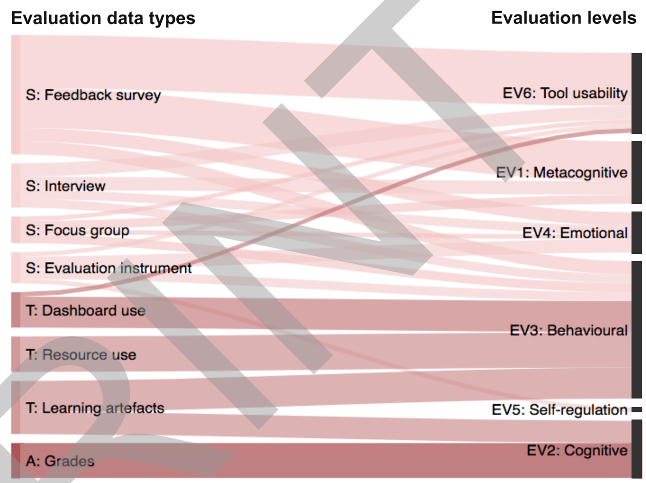

Figure 2: The evaluation data types in relation to the criteria assessed in the evaluation of the dashboards. Self-reported data (S) is illustrated with light red, tracked data (T) with red and assessment data (A) with dark red.

Figure 2 illustrates the types of data used to operationalise the different evaluation criteria listed in Table 1. The graph shows that for assessing aspects on the metacognitive evaluation level, researchers used only self-reported data which was collected through feedback surveys (9 papers), interviews (3) and focus groups (2). The effect of the dashboard on learners' cognitive competence was operationalised through learners' performance. This was measured through assessment data, i.e., grades and learning gain (8 papers) and the quality of learning artefacts, i.e., discussion posts or effectiveness scores (5 papers). Impact on behaviour was mostly investigated by computing engagement metrics from trace log data (14 papers). We also identified instances where researchers used surveys (3 papers), interviews (2) or focus groups (2) to ask learners to interpret data displayed on dashboards and make recommendations of corrective actions. Other works used evaluation instruments to measure learners' time-management skills [57] and approaches to learning [29]. Self-reported data were used for evaluating aspects of the emotional level. Impact on learner motivation was measured through self-reported data collected through surveys (2), interviews (1), focus groups (1) and evaluation instruments (3), while the impact on affect was measured in interviews (1) and by analysing the transitions between affective states in self-reported data (1). The tool's usability was mostly evaluated through self-reports. These data were collected through feedback surveys (12), interviews (3) and focus groups (1) conducted at the end of a dashboard's evaluation phase. Only one paper [57] used the System Usability Scale

Table 4: The dashboard evaluation criteria assessed in the 26 papers included in this literature review and the data types used to judge each criterion.

		Self-reported data				Tracked data			Assessment data
		Feedback survey	Interviews	Focus group	Evaluation instruments	Resources use	Learning artefacts	Dashboard use	Grades
EV1: Metcognitive	Understanding	[4, 15, 18, 28, 37] [42, 50, 51]	[16, 50, 63]	[34]	-	-	-	-	-
	Agreement	[37, 42]	-	[34]	-	-	-	-	-
	Impact on awareness and reflection	[15, 18, 37, 38, 42] [50, 51]	[16, 50]	[34, 58]	-	-	-	-	-
EV2: Cognitive	Impact on performance	-	-	-	-	-	[7, 19, 28, 29, 38]	-	[19, 28, 31, 36] [50, 51, 54, 57]
EV3: Behavioural	Impact on behaviour	[4, 18, 50]	[16, 63]	[34, 58]	[29, 57]	[18, 19, 28, 37] [38, 51, 57, 63]	[7, 19, 28, 29, 38] [51, 63]	-	-
	Usage of the system	-	[63]	-	-	-	-	[13, 28, 31, 38] [44, 50, 57]	-
EV4: Emotional	Impact on motivation	[28, 38]	[16]	[58]	[7, 16, 28]	-	-	-	-
	Impact on affect	[44]	[63]	-	-	-	-	-	-
EV5: Self-regulation	SRL	-	-	-	[57]	-	-	-	-
EV6: Tool usability	Satisfaction	[31, 38, 50]	[50]	-	-	-	-	-	-
	Usability	[3, 12, 28, 31, 38] [50, 58]	[50, 63]	-	[57]	-	-	[28]	-
	Usefulness	[12, 13, 26, 28, 31] [35, 37, 38, 42, 50] [58]	[16, 63]	[34]	-	-	-	-	-

[11], a standardised tool for measuring tool usability. In one particular case, Guerra et al. [28] used logs of learners' interactions with the dashboard to evaluate the complexity of the interface and its usability. Table 4 presents a detailed overview on what data was used for operationalising the different evaluation criteria.

3.3 Learners' evaluation of reference frames

According to the framework for designing pedagogical interventions to support student use of learning analytics proposed by Wise [64], learners need a "representative reference frame" for interpreting their data on a dashboard. In our previous work [30], we identified three types of reference frames: (i) social, i.e., comparison with peers, (ii) achievement, i.e., distance towards goals, and (iii) progress, i.e., comparison with an earlier self. Now, we looked at how these different frames of references were perceived by learners.

Two papers reported what reference frames students preferred. Surveying students about their preferred features of the dashboard, Konert et al. [35] found that most respondents liked the comparison to other peers' knowledge and time investment. In contrast to that, Tabuenca et al. [57] found that students preferred personal analytics to social analytics and teacher's estimations. A possible explanation for these contrasting results might be the fact that on the PeerLA dashboard [35] learners could choose to compare themselves against students that had similar goals to theirs.

Looking at the effects of social comparison on behaviour and performance, Guerra et al. [28] evaluated how much the users used the social comparison elements and discovered that those who clicked on the social features had more activities on problems and examples, but they did not imply any causal link. Ruiz et al. [50] also concluded that tracking, visualising and comparing emotions to the group helped students to increase attention or effort while studying, improving students' general mood. Similarly, Davis et al. [19] presented evidence that a social comparison feedback system that promoted learners' awareness of their successful classmates' behaviour lead to statistically significant increases in the course completion rate in MOOCs.

Several papers investigated in more detail how learners perceive different frames of reference. Davis et al. [19] reported that the benefits of their feedback system were limited to learners who are highly educated. Kim et al. [31] also found that the information presented on the dashboard was perceived differently by learners depending on their academic achievement level. Low achievers who did not know their learning status compared to others, found the dashboard motivating, while high achievers are already strongly motivated and comparison with peers might not be as motivating as seeing their achievement reported to their own set goals. Corrin et al. [17] observed the same effect: being compared with the class average had a positive influence on most students' motivation, but it also distracted half of the students from their performance goal. Once learners saw that they were performing slightly above the average, they would feel satisfied and not strive to reach their previously set goal. The learners that were not motivated by the dashboard were the ones that were above the class average or were meeting their performance goals on the subject. Contrasting results were obtained by Tan et al. [58]: highly-performing students were even more motivated by the "healthy peer pressure" and the "informal competition" that drove them to engage with learning in a more stimulating way. On the other hand, learners whose performance was on the outside (of the social network graph) or at the bottom (of the class) felt demoralised and stressed. These students preferred when their learning performance was contextualised criterion-based, i.e., according to learning goals, and self-referenced, i.e., according to their previous performance, rather than norm-referenced, i.e., according to the performance of their peers. Similar results were also found in [63], the comparison with averages of their peers was found interesting, motivating and helpful by some, but also surprising, intimidating and stressful, especially by low performing students that were below average.

4 DISCUSSION

Through this literature review, we build on our previous work to further investigate the relation between learning sciences and learning analytics. Here, we looked at how learning analytics dashboards for learners were evaluated and how educational concepts borrowed from the learning sciences were taken into consideration when evaluating these dashboards. The discussion section ends with a summary of recommendations for dashboard design (D) and evaluation (E) that are referenced through the whole section, e.g., D1, E2.

4.1 Shortcomings of dashboard evaluations

According to Park & Jo [46], a visualisation tool is meaningful only when it generates the intended changes in the users. The goal of the dashboard articulates these intended changes. In our previous work [30], we classified the goals of 26 dashboards based on the competence they aimed to affect in learners: metacognitive, cognitive, behavioural, emotional and self-regulation. In the present study, we looked at whether changes in these competences were evaluated and, if yes, how.

A first insight is that less than one-third of the analysed papers targeted the cognitive competence. Although awareness and reflection are the main goals for many dashboards [30], reflection should not be seen as an end in itself, but rather as a mechanism that can improve teaching and thus maximise learning [41]. LA dashboards should be designed and evaluated as pedagogical tools which catalyse changes also in the cognitive, behavioural and emotional competencies, and not only on the metacognitive level (D1).

On the other hand, learners' cognitive competence was evaluated in almost twice as many papers as in which it was targeted as a competence. Similarly, the behavioural level was evaluated in three-quarters of the analysed papers, although less than half of the papers indicated their dashboards supported behavioural changes. We can attribute this inconsistency to the fact that learners' "observable behaviour" [45] is more easily extracted from log traces that are automatically recorded by the platform in which the dashboard is embedded rather than collecting subjective evaluations from users. Indeed, the main source of data, in this case, was tracked data. This finding suggests that similar to the dashboard design [30], the evaluation is also driven by the availability of data rather than a clear pedagogical focus (E1).

Secondly, we were surprised to find out that evaluating a dashboard's acceptance, usefulness and ease-of-use as perceived by learners is central in many of the analysed papers. In almost all cases, dashboard designers used feedback questionnaires and interviews for confirmation that learners are satisfied and find the visualisations useful. This is not surprising considering that several authors highlighted perceived ease-of-use as a significant factor that influences the adoption of learning analytics tools [2, 31]. Nonetheless, the surveys used in the evaluation of the dashboards asked learners to rate the usability and usefulness of learning analytics tools and the results were analysed in a quantitative manner. In order to bring insight into what is and what is not useful for learners, surveys could include open questions or could be followed up by interviews or focus groups. However, although we agree that tool acceptance is essential, this should be a secondary focus of the evaluation, and the primary focus should be whether the dashboard brings any benefit to learners (E1).

We found very few papers that assessed whether dashboards have any effect on the emotional competences. Learners' positive and negative emotions have been shown to have a big influence on online learning behaviour [59] and certain features of learning analytics dashboards can lead to feelings of disappointment, intimidation or stress [58, 63]. We hypothesise that evaluating changes in learners' affect and motivation when using dashboards would lead to more effective and accommodating dashboards, contributing to solving the low uptake problem [39] (E1).

Another observation relates to the evaluation of changes on the metacognitive level. Half of the papers we analysed requested feedback from learners to assess if learners understood the data displayed on the dashboard or if the dashboard increased their awareness and offered any support for reflection. Supporting the efforts of the community in evaluating the effects of learning analytics tools, Scheffel et al. [52] devised a learning analytics evaluation framework (EFLA) that measures aspects of the metacognitive competence: understanding of data, awareness, reflection and impact on the learner, through an 8-item questionnaire. In remarkably few cases, designers asked learners whether they agree or disagree with the information presented on the dashboard. Considering that building trust and confidence in learning analytics tools is a major concern for learning analytics [27], assessing the accuracy of the information presented on dashboards from the learner's point of view, however, should be a priority when it comes to evaluating learning analytics dashboards (E2).

Finally, a considerable number of papers evaluated the dashboards on more than one level and using data from multiple sources. Several of these dashboard studies performed ingenious evaluations using different types of data. For example, Muldner et al. [44] explored the interplay between usage of the tool and impact on affect, while [63] used interviews and data about the resource use and learning artefacts to identify differences between intended behaviour and actual behaviour. Loboda et al. [38] analysed correlations between access to material through the dashboard and final grades, while the tool usage data was used to evaluate whether accessing and using the tool had any effect on learners' performance. Hatala et al. [29] and Beheshitha et al. [7] investigated the usage of learning dashboards in relation to the quantity and quality of the discussion posts while considering learners' self-reported approaches to learning and achievement goal orientation. These examples showcase diverse approaches to evaluating the impact of dashboards on learners and the possibilities collected data can create for analyses. Therefore, in order to reliably evaluate LA tools, we recommend that researchers validate subjective tool evaluations from learners, e.g., from feedback surveys, interviews and focus groups, with objective data extracted from trace logs and assessment data in order to answer the question whether the dashboard effect perceived by learners can also be observed in their interaction with the online learning environment and whether the learning outcomes have improved (E3).

4.2 Educational concepts used in dashboard evaluation

We identified seven papers that take into account theories both in the design and the evaluation of the dashboards. The theories and concepts considered in the evaluation of the dashboard are self-regulated learning (1 paper), achievement goal orientation (2), social comparison (2), deep and surface approaches to learning (1) and learning power (1).

In our previous work, we showed that self-regulated learning is the most common educational concept used in the design of learning analytics dashboards and many dashboards are built with the goal of supporting self-regulated learning [30]. However, only one of the analysed papers evaluated improvement in self-regulated skills. Tabuenca et al. [57] used a dedicated evaluation instrument to measure the level of SRL skill several times throughout the course. Other papers evaluated specific SRL skills: help-seeking behaviour in [18] and the ability to plan new or amend study strategies in [16]. As the dashboards that relied on self-regulated learning theories focused more on supporting the self-evaluation step, rather than the planning or monitoring, the dashboards were usually evaluated by asking learners if they found the visualisations of the dashboards helpful in supporting awareness and reflection. This analysis was occasionally complemented by exploring trace logs to identify changes in behaviour or improvements in final grades. Similarly, the influence of social comparison, achievement goal orientation and approaches to learning were studied either through learner surveys or by checking for changes in behaviour and performance. However, since all papers included in this review considered educational concepts borrowed from learning sciences or psychology in their designs, we expected more papers to take them into account also in the evaluation phase, especially in interpreting the results (E4).

4.3 Evaluation of reference frames

Comparison with peers was the most common reference frame implemented in dashboards in order to offer learners an anchor point for self-evaluation [30]. Comparison with peers was usually used to motivate students to work harder and increase their engagement, sometimes by "inducing a feeling of being connected with and supported by their peers" [60]. Among the concepts mentioned by the analysed papers, there were only two theories that would justify the use of comparison with peers: social comparison theory and achievement goal orientation theory. Social comparison [23] states that we establish our self-worth by comparing ourselves to others when there are no objective means of comparison, while achievement goal orientation theory distinguishes between the different motivations why one engages in an achievement task [21]. Learners who are motivated by mastery goals focus on acquiring knowledge and mastering the tasks, while learners who have performance goals are motivated when demonstrating their ability and measuring their skill in comparison to others. However, few papers motivated the use of comparison with peers.

As we looked into how learners perceived their comparison with peers, we have found contrasting results. While some authors argue that top-achieving students were motivated by seeing their success in comparison to others, other authors found evidence that being at the top of the class obscured learners' initial learning goals. At the same time, comparison with peers has been shown to bring feelings of distress, demotivation and disappointment, especially in low-performing students. Individual differences between learners, like the ones described through the achievement goal orientation theory, could explain these contrasting results. Beheshitha et al. [7] have indeed presented initial results that after controlling for achievement goals, some learning analytics visualisations had positive and some had negative effects on students' quantity and quality of discussion posts evaluated through discourse features. However, further research has to be invested into understanding and explaining learners' different responses to different reference frames, particularly the often-present comparison with peers (D3).

These results show that dashboard evaluations rarely consider concepts from learning sciences. Even more importantly, these concepts are not consulted in order to explain different learning behaviours, or reactions of learners when using the dashboards (D4). Furthermore, we encountered a low number of validated instruments that were used to assess either the learners' skills or the tools, suggesting a lack of consistency across the field regarding how learning analytics systems are evaluated. Our findings imply that overlooking learning sciences research when designing and evaluating learning analytics dashboards can result in inadequate tools that are quickly dismissed by learners (E5). Linking the fields of learning analytics and learning sciences could bring valuable additions in terms of how evaluation is conducted and how the effects of dashboards are measured, bringing learning dashboards closer to wide-spread adoption.

4.4 Recommendations for dashboard design and evaluation

We have compiled a set of recommendations for the design (D) and evaluation (E) of learning analytics dashboards for learners that summarise the insights gathered through our literature reviews published in this work as well as in Jivet et al. [30]. Our insights complement the open issues of learning dashboard research identified by Schwendimann et al. [53] and the recommendations for practice and future research outlined by Bodily & Verbert [9], offering suggestions for integrating learning sciences research into the development of learning analytics tools.

- D1 LA dashboards should be designed as pedagogical tools that enhance awareness and reflection as a means to catalyse changes in the cognitive, behavioural and emotional competences.
- D2 Educational concepts from learning sciences should be used to motivate design decisions.
- D3 Comparison with peers should be used cautiously.
- D4 Do not assume the dashboard will have the same effect on all its users, but rather seek to determine which group of learners benefit the most and how to customise the dashboard to provide the same support to all its users.
- D5 The dashboard should be seamlessly integrated into the online learning environment and into the usual learning activities of the learner.
- E1 Dashboard evaluation should focus (primarily) on whether its goals are fulfilled, (secondarily) on the impact on learners'

affect and motivation, and (finally) on the usability of the tool.

E2 The evaluation of a tool's usability and usefulness should not be limited to whether users find the tool usable and useful, but in order to build trust and confidence in learning analytics tools, it should also assess whether learners understand the data, how much they agree with it and how they interpret it.

E3 Dashboard evaluation should use data triangulation to validate its effects with self-reported data, tracked data as well as assessment data.

E4 The evaluation should include an assessment of the design features that rely on educational concepts.

E5 Validated measurement instruments should be used to assess if the dashboard had any impact on the learner or if learner characteristics as measured through these instruments play a role in how learners perceive the dashboard and how they respond to it.

5 CONCLUSION

Throughout this literature review, we looked at the use of educational concepts in learning analytics dashboards for learners. As we put particular focus on the evaluation of these dashboards, we have observed that there is a strong mismatch between the goal of the dashboard and its evaluation. The majority of dashboards aim to support the metacognitive level and a very low number seeks to support learners cognitively or emotionally. However, most dashboard evaluations focus on assessing the tool's usability and the impact on the behavioural competence. The effects on the cognitive and emotional levels received very little attention overall. This finding strengthens our conclusion from the previous study that the development of learning analytics dashboards is still driven by the need to leverage the learning data available, rather than a clear pedagogical focus of improving and supporting learning.

Secondly, the preponderance of self-reported data among the types of data collected for dashboard evaluation suggests that the potential of leveraging the "analytics" was not used as much in the evaluation as in the design and implementation of dashboards. Tool acceptance and usefulness as perceived by learners are indeed important factors that influence the adoption of learning analytics tools. However, complementing the feedback gathered through self-reports with insights gained from exploring the learning data collected from trace logs will provide more credibility to the results, ensuring both researchers and practitioners of the benefit for learning brought by learning analytics dashboards.

Finally, through our analyses, we identified a gap between learning sciences and learning analytics. We propose three approaches for using research from learning sciences in developing and evaluating learning analytics dashboards. Firstly, learning analytics researchers can support decisions related to the dashboard design through educational concepts. Additionally, they can use existing validated measurement instruments for assessing whether the tools support learners' competences. Finally, learning theories and other concepts can be used as a lens for understanding contrasting results when evaluating learners' responses to using learning analytics dashboards.

REFERENCES

[1] Sema Alay and Settar Kocak. 2002. Validity and reliability of time management questionnaire. *Hacettepe Universitesi Egitim FakuÌĹltesi Derg* 22 (2002), 9–13.
[2] Liaqat Ali, Mohsen Asadi, Dragan Gašević, Jelena Jovanović, and Marek Hatala. 2013. Factors influencing beliefs for adoption of a learning analytics tool: An empirical study. *Computers & Education* 62 (2013), 130–148.
[3] Naif Radi Aljohani and Hugh C Davis. 2013. Learning analytics and formative assessment to provide immediate detailed feedback using a student centered mobile dashboard. In *Next Generation Mobile Apps, Services and Technologies (NGMAST), 2013 Seventh International Conference on*. IEEE, 262–267.
[4] Antonio R Anaya, Manuel Luque, and Manuel Peinado. 2016. A visual recommender tool in a collaborative learning experience. *Expert Systems with Applications* 45 (2016), 248–259.
[5] Ryan S Baker. 2016. Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education* 26, 2 (2016), 600–614.
[6] Lucy Barnard, William Y Lan, Yen M To, Valerie Osland Paton, and Shu-Ling Lai. 2009. Measuring self-regulation in online and blended learning environments. *The Internet and Higher Education* 12, 1 (2009), 1–6.
[7] Sanam Shirazi Beheshitha, Marek Hatala, Dragan Gašević, and Srećko Joksimović. 2016. The role of achievement goal orientations when studying effect of learning analytics visualizations. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. ACM, 54–63.
[8] John Biggs, David Kember, and Doris YP Leung. 2001. The revised two-factor study process questionnaire: R-SPQ-2F. *British journal of educational psychology* 71, 1 (2001), 133–149.
[9] Robert Bodily and Katrien Verbert. 2017. Review of research on student-facing learning analytics dashboards and educational recommender systems. *IEEE Transactions on Learning Technologies* (2017).
[10] Robert Bodily and Katrien Verbert. 2017. Trends and issues in student-facing learning analytics reporting systems research. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. ACM, 309–318.
[11] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.
[12] Susan Bull, Blandine Ginon, Clelia Boscolo, and Matthew Johnson. 2016. Introduction of learning visualisations and metacognitive support in a persuadable open learner model. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. ACM, 30–39.
[13] Susan Bull, Matthew D Johnson, Carrie Demmans Epp, Drew Masci, Mohammad Alotaibi, and Sylvie Girard. 2014. Formative assessment and meaningful learning analytics. In *Advanced Learning Technologies (ICALT), 2014 IEEE 14th International Conference on*. IEEE, 327–329.
[14] Sven Charleer, Joris Klerkx, Erik Duval, Tinne De Laet, and Katrien Verbert. 2016. Creating effective learning analytics dashboards: Lessons learnt. In *European Conference on Technology Enhanced Learning*. Springer, 42–56.
[15] Chih-Yueh Chou, Shu-Fen Tseng, Wen-Chieh Chih, Zhi-Hong Chen, Po-Yao Chao, K Robert Lai, Chien-Lung Chan, Liang-Chih Yu, and Yi-Lung Lin. 2015. Open Student Models of Core Competencies at the Curriculum Level: Using Learning Analytics for Student Reflection. *IEEE Transactions on Emerging Topics in Computing* (2015).
[16] Linda Corrin and Paula de Barba. 2014. Exploring students' interpretation of feedback delivered through learning analytics dashboards. In *Proceedings of the ascilite 2014 conference*. 629–633.
[17] Linda Corrin and Paula de Barba. 2015. How do students interpret feedback delivered via dashboards?. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*. ACM, 430–431.
[18] Samantha G Daley, Garron Hillaire, and LeeAnn M Sutherland. 2016. Beyond performance data: Improving student help seeking by collecting and displaying influential data in an online middle-school science curriculum. *British Journal of Educational Technology* 47, 1 (2016), 121–134.
[19] Dan Davis, Ioana Jivet, René F Kizilcec, Guanliang Chen, Claudia Hauff, and Geert-Jan Houben. 2017. Follow the successful crowd: raising MOOC completion rates through social comparison at scale. In *Proceedings of the 7th International Conference on Learning Analytics and Knowledge*.
[20] Eva Durall and Begoña Gros. 2014. Learning Analytics as a Metacognitive Tool.. In *CSEDU (1)*. 380–384.
[21] Andrew J Elliot, Kou Murayama, and Reinhard Pekrun. 2011. A 3× 2 achievement goal model. *Journal of Educational Psychology* 103, 3 (2011), 632.
[22] Rebecca Ferguson. 2012. Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning* 4, 5-6 (2012), 304–317.
[23] Leon Festinger. 1954. A theory of social comparison processes. *Human relations* 7, 2 (1954), 117–140.
[24] Stephen Few. 2013. *Information Dashboard Design: Displaying data for at-a-glance monitoring*. Analytics Press.
[25] Dragan Gašević, Shane Dawson, and George Siemens. 2015. Let's not forget: Learning analytics are about learning. *TechTrends* 59, 1 (2015), 64–71.
[26] Sten Govaerts, Katrien Verbert, and Erik Duval. 2011. Evaluating the student activity meter: two case studies. In *International Conference on Web-Based Learning*.

Springer, 188–197.
[27] Wolfgang Greller and Hendrik Drachsler. 2012. Translating learning into numbers: A generic framework for learning analytics. *Educational technology & society* 15, 3 (2012), 42–57.
[28] Julio Guerra, Roya Hosseini, Sibel Somyurek, and Peter Brusilovsky. 2016. An intelligent interface for learning content: Combining an open learner model and social comparison to support self-regulated learning and engagement. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*. ACM, 152–163.
[29] Marek Hatala, Sanam Shirazi Beheshitha, and Dragan Gasevic. 2016. Associations Between Students' Approaches to Learning and Learning Analytics Visualizations.. In *LAL@ LAK*. 3–10.
[30] Ioana Jivet, Maren Scheffel, Hendrik Drachsler, and Marcus Specht. 2017. Awareness is not enough: pitfalls of learning analytics dashboards in the educational practice. In *European Conference on Technology Enhanced Learning*. Springer, 82–96.
[31] Jeonghyun Kim, Il-Hyun Jo, and Yeonjeong Park. 2016. Effects of learning analytics dashboard: analyzing the relations among dashboard utilization, satisfaction, and learning achievement. *Asia Pacific Education Review* 17, 1 (2016), 13–24.
[32] Donald L Kirkpatrick. 1977. Evaluating training programs: Evidence vs. proof. *Training Dev J* (1977).
[33] Joris Klerkx, Katrien Verbert, and Erik Duval. 2014. Enhancing learning with visualization techniques. In *Handbook of research on educational communications and technology*. Springer, 791–807.
[34] Elizabeth Koh, Antonette Shibani, Jennifer Pei-Ling Tan, and Helen Hong. 2016. A pedagogical framework for learning analytics in collaborative inquiry tasks: an example from a teamwork competency awareness program. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. ACM, 74–83.
[35] Johannes Konert, Christoph Bohr, Henrik Bellhäuser, and Christoph Rensing. 2016. PeerLA-Assistant for individual learning goals and self-regulation competency improvement in online learning scenarios. In *Advanced Learning Technologies (ICALT), 2016 IEEE 16th International Conference on*. IEEE, 52–56.
[36] Vivekanandan Kumar, David Boulanger, Jeremie Seanosky, Karthikeyan Panneerselvam, Thamarai Selvi Somasundaram, et al. 2014. Competence analytics. *Journal of Computers in Education* 1, 4 (2014), 251–270.
[37] Ming Liu, Rafael A Calvo, Abelardo Pardo, and Andrew Martin. 2015. Measuring and visualizing students' behavioral engagement in writing activities. *IEEE Transactions on learning technologies* 8, 2 (2015), 215–224.
[38] Tomasz D Loboda, Julio Guerra, Roya Hosseini, and Peter Brusilovsky. 2014. Mastery grids: An open source social educational progress visualization. In *European Conference on Technology Enhanced Learning*. Springer, 235–248.
[39] Steven Lonn, Stephen J Aguilar, and Stephanie D Teasley. 2015. Investigating student motivation in the context of a learning analytics intervention during a summer bridge program. *Computers in Human Behavior* 47 (2015), 90–97.
[40] Zahia Marzouk, Mladen Rakovic, Amna Liaqat, Jovita Vytasek, Donya Samadi, Jason Stewart-Alonso, Ilana Ram, Sonya Woloshen, Philip H Winne, and John C Nesbit. 2016. What if learning analytics were based on learning science? *Australasian Journal of Educational Technology* 32, 6 (2016).
[41] Lynn McAlpine and Cynthia Weston. 2000. Reflection: Issues related to improving professors' teaching and students' learning. *Instructional Science* 28, 5 (2000), 363–385.
[42] Carolina Mejia, Beatriz Florian, Ravi Vatrapu, Susan Bull, Sergio Gomez, and Ramon Fabregat. 2017. A novel web-based approach for visualization and inspection of reading difficulties on university students. *IEEE Transactions on Learning Technologies* 10, 1 (2017), 53–67.
[43] David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G Altman, Prisma Group, et al. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS med* 6, 7 (2009), e1000097.
[44] Kasia Muldner, Michael Wixon, Dovan Rai, Winslow Burleson, Beverly Woolf, and Ivon Arroyo. 2015. Exploring the impact of a learning dashboard on student affect. In *International Conference on Artificial Intelligence in Education*. Springer, 307–317.
[45] Kyparisia A Papanikolaou. 2015. Constructing interpretative views of learners' interaction behavior in an open learner model. *IEEE Transactions on Learning Technologies* 8, 2 (2015), 201–214.
[46] Yeonjeong Park and Il-Hyun Jo. 2015. Development of the Learning Analytics Dashboard to Support Students' Learning Performance. *J. UCS* 21, 1 (2015), 110–133.
[47] Paul R Pintrich. 2000. Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of educational psychology* 92, 3 (2000), 544.
[48] Paul R Pintrich et al. 1991. A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ). (1991).
[49] José A Ruipérez-Valiente, Pedro J Muñoz-Merino, Derick Leony, and Carlos Delgado Kloos. 2015. ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. *Computers in Human Behavior* 47 (2015), 139–148.
[50] Samara Ruiz, Sven Charleer, Maite Urretavizcaya, Joris Klerkx, Isabel Fernández-Castro, and Erik Duval. 2016. Supporting learning by considering emotions: tracking and visualization a case study. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. ACM, 254–263.
[51] Maren Scheffel, Hendrik Drachsler, Karel Kreijns, Joop de Kraker, and Marcus Specht. 2017. Widget, widget as you lead, I am performing well indeed!: using results from an exploratory offline study to inform an empirical online study about a learning analytics widget in a collaborative learning environment. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. ACM, 289–298.
[52] Maren Scheffel, Hendrik Drachsler, Christian Toisoul, Stefaan Ternier, and Marcus Specht. 2017. The proof of the pudding: examining validity and reliability of the evaluation framework for learning analytics. In *European Conference on Technology Enhanced Learning*. Springer, 194–208.
[53] Beat Schwendimann, Maria Rodriguez-Triana, Andrii Vozniuk, Luis Prieto, Mina Boroujeni, Adrian Holzer, Denis Gillet, and Pierre Dillenbourg. 2016. Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies* (2016).
[54] Simon Buckingham Shum and Ruth Deakin Crick. 2012. Learning dispositions and transferable competencies: pedagogy, modelling and learning analytics. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. ACM, 92–101.
[55] George Siemens and Dragan Gasevic. 2012. Guest editorial-Learning and knowledge analytics. *Educational Technology & Society* 15, 3 (2012), 1–2.
[56] Dan Suthers and Katrien Verbert. 2013. Learning analytics as a middle space. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*. ACM, 1–4.
[57] Bernardo Tabuenca, Marco Kalz, Hendrik Drachsler, and Marcus Specht. 2015. Time will tell: The role of mobile learning analytics in self-regulated learning. *Computers & Education* 89 (2015), 53–74.
[58] Jennifer Pei-Ling Tan, Simon Yang, Elizabeth Koh, and Christin Jonathan. 2016. Fostering 21st century literacies through a collaborative critical reading and learning analytics environment: user-perceived benefits and problematics. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. ACM, 430–434.
[59] Dirk T Tempelaar, Alexandra Niculescu, Bart Rienties, Wim H Gijselaers, and Bas Giesbers. 2012. How achievement emotions impact students' decisions for online learning, and what precedes those emotions. *The Internet and Higher Education* 15, 3 (2012), 161–169.
[60] Rémi Venant, Philippe Vidal, and Julien Broisin. 2016. Evaluation of Learner Performance During Practical Activities: An Experimentation in Computer Education. In *Advanced Learning Technologies (ICALT), 2016 IEEE 16th International Conference on*. IEEE, 237–241.
[61] Katrien Verbert, Erik Duval, Joris Klerkx, Sten Govaerts, and José Luis Santos. 2013. Learning analytics dashboard applications. *American Behavioral Scientist* 57, 10 (2013), 1500–1509.
[62] Katrien Verbert, Sten Govaerts, Erik Duval, Jose Luis Santos, Frans Van Assche, Gonzalo Parra, and Joris Klerkx. 2014. Learning dashboards: an overview and future research opportunities. *Personal and Ubiquitous Computing* 18, 6 (2014), 1499–1514.
[63] Alyssa Wise, Yuting Zhao, and Simone Hausknecht. 2014. Learning analytics for online discussions: Embedded and extracted approaches. *Journal of Learning Analytics* 1, 2 (2014), 48–71.
[64] Alyssa Friend Wise. 2014. Designing pedagogical interventions to support student use of learning analytics. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*. ACM, 203–211.
[65] W-H Wu, H-C Hsiao, P-L Wu, C-H Lin, and S-H Huang. 2012. Investigating the learning-theory foundations of game-based learning: a meta-analysis. *Journal of Computer Assisted Learning* 28, 3 (2012), 265–279.
[66] Yesom Yoo, Hyeyun Lee, Il-Hyun Jo, and Yeonjeong Park. 2015. Educational dashboards for smart learning: Review of case studies. In *Emerging Issues in Smart Learning*. Springer, 145–155.
[67] Barry J Zimmerman. 1990. Self-regulated learning and academic achievement: An overview. *Educational psychologist* 25, 1 (1990), 3–17.
[68] Barry J Zimmerman, M Boekarts, PR Pintrich, and M Zeidner. 2000. A social cognitive perspective. *Handbook of self-regulation* 13, 1 (2000), 695–716.